

UNITED STATES PATENT APPLICATION
FOR

METHOD AND SYSTEM TO INSTRUMENT
VIRTUAL FUNCTION CALLS

Inventors:
BRIAN FAHS
ROBERT HUNDT
TARA KRISHNASWAMY

Prepared by:
WAGNER, MURABITO & HAO LLP
Two North Market Street
Third Floor
San Jose, California 95113

5

METHOD AND SYSTEM TO INSTRUMENT VIRTUAL FUNCTION CALLS

TECHNICAL FIELD

10 The present claimed invention relates to instrumentation of a computer program. More specifically, the present claimed invention relates to instrumenting of virtual functions.

BACKGROUND ART

15 Over recent years, the computing community developed a strong set of tools and methods used to analyze and monitor run-time behavior of a program. One type of performance analysis is referred to as instrumentation. Measurements such as basic-block coverage and function invocation counting can be accurately made using
20 instrumentation. One specific type of code instrumentation is referred to as dynamic binary instrumentation. Dynamic binary instrumentation allows program instructions to be changed on-the-fly. Additionally, dynamic binary instrumentation, as opposed to static instrumentation, is performed at run-time of a program and only instruments those parts of
25 an executable that are actually executed. This minimizes the overhead imposed by the instrumentation process itself. Furthermore, performance analysis tools based on dynamic binary instrumentation require no special preparation of an executable such as, for example, a modified build or link process.

30

Certain modern programming languages such as, for example, C++ offer the ability to inherit so called derived objects from other base objects. This concept is commonly known as inheritance in the object oriented programming domain. Often, these base and/or derived objects
35 use what are known as virtual functions. Hence, it is possible in certain instances to make a call to a virtual function. To accomplish this, the compiler generates an array of function pointers, known as a virtual table, for each object type that contains at least one virtual function. During the virtual function call, this virtual table is indexed to obtain a
40 function pointer, and then an indirect call is made using that function

pointer. Such tables must be created because the actual function call made may not be determinable at compile time. Additionally, it is not possible, at present, to readily instrument or analyze such virtual function calls.

5

Furthermore, programmers are often interested in deciphering the type of call which is made to a function such as, for example, a virtual function. More specifically, various function call types such as, for example, direct function calls, indirect function calls, and virtual

10 function calls have differing costs associated therewith. That is, the various function calls differ in terms of the cycles and/or instructions executed in performing the function call. Unfortunately, at present, there is no current method for determining the call type for a virtual function.

15

Thus, a need has arisen for a method and system for analyzing a virtual function including determining the type of call made to the virtual function.

DISCLOSURE OF THE INVENTION

The present invention provides a method and system for method and system for instrumenting a virtual function including determining the type of call made to the virtual function.

5

Specifically, in one method embodiment, the present invention determines a call type for a virtual function. The present embodiment then locates a virtual table corresponding to a virtual function and replaces an existing address for the virtual function with a new address
10 for the virtual function in the virtual table. In this embodiment, the new address points to instrumentation code. Upon a call to the virtual function, the present embodiment then loads the new address from the virtual table such that execution is directed to the instrumentation code. The present embodiment continues execution and executes the
15 instrumentation code and delivers control to an instrumenting application.

These and other technical advantages of the present invention will no doubt become obvious to those of ordinary skill in the art after having
20 read the following detailed description of the preferred embodiments which are illustrated in the various drawing figures.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and form a part of this specification, illustrate embodiments of the invention and, together with the description, serve to explain the principles of the invention:

FIGURE 1 is a schematic diagram of an exemplary computer system used to perform steps of the present method in accordance with various embodiments of the present claimed invention.

FIGURE 2 is a flow chart of steps performed to determine a call type of a virtual function in accordance with one embodiment of the present claimed invention.

FIGURE 3 is a flow chart of steps performed in performing instrumentation of a virtual function in accordance with one embodiment of the present claimed invention.

FIGURE 4 is a flow chart of steps performed in instrumentation of a virtual function by readdressing a virtual table in accordance with one embodiment of the present claimed invention.

FIGURE 5 is a flow chart of steps performed in instrumentation of a virtual function by readdressing of a virtual table and controlling execution via an instrumentor in accordance with one embodiment of the present claimed invention.

FIGURE 6 is a flow chart of steps performed in instrumentation of a virtual function by readdressing of a virtual table and controlling execution via a target process in accordance with one embodiment of the present claimed invention.

FIGURE 7 is a flow chart of steps performed to determine a call type of a virtual function and to determine from which the virtual function has been called in accordance with one embodiment of the present claimed invention.

FIGURE 8 is a flow chart of steps performed in instrumentation of a virtual function by readdressing a virtual table and maintaining a

mapping of virtual tables addresses in accordance with one embodiment of the present claimed invention.

The drawings referred to in this description should be understood
5 as not being drawn to scale except if specifically noted.

BEST MODES FOR CARRYING OUT THE INVENTION

Reference will now be made in detail to the preferred embodiments of the invention, examples of which are illustrated in the accompanying drawings. While the invention will be described in conjunction with the preferred embodiments, it will be understood that they are not intended to limit the invention to these embodiments. On the contrary, the invention is intended to cover alternatives, modifications and equivalents, which may be included within the spirit and scope of the invention as defined by the appended claims. Furthermore, in the following detailed description of the present invention, numerous specific details are set forth in order to provide a thorough understanding of the present invention. However, it will be obvious to one of ordinary skill in the art that the present invention may be practiced without these specific details. In other instances, well known methods, procedures, components, and circuits have not been described in detail as not to unnecessarily obscure aspects of the present invention.

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussions, it is appreciated that throughout the present invention, discussions utilizing terms such as "determining", "instrumenting", "overwriting", "executing", "performing", or the like, refer to the actions and processes of a computer system, or similar electronic computing device. The computer system or similar electronic computing device manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission, or display devices. The present invention is also well suited to the use of other computer systems such as, for example, optical and mechanical computers.

COMPUTER SYSTEM ENVIRONMENT OF THE PRESENT INVENTION

With reference now to Figure 1, portions of the present method and system are comprised of computer-readable and computer-executable instructions which reside, for example, in computer-usable media of a computer system. Figure 1 illustrates an exemplary computer system 100

used in accordance with one embodiment of the present invention. It is appreciated that system 100 of Figure 1 is exemplary only and that the present invention can operate on or within a number of different computer systems including general purpose networked computer systems, embedded computer systems, routers, switches, server devices, client devices, various intermediate devices/nodes, stand alone computer systems, and the like. Additionally, computer system 100 of Figure 1 is well adapted having computer readable media such as, for example, a floppy disk, a compact disc, and the like coupled thereto. Such computer readable media is not shown coupled to computer system 100 in Figure 1 for purposes of clarity. Additionally, portions of the present embodiment are well suited to operating in conjunction with various mobile clients such as, for example, a cell phone, personal digital assistant (PDA), laptop computer, pager, and the like.

System 100 of Figure 1 includes an address/data bus 102 for communicating information, and a central processor unit 104 coupled to bus 102 for processing information and instructions. As an example, central processor unit 104 may be an IA-64 microprocessor architecture by Intel Corporation of Santa Clara, California. System 100 also includes data storage features such as a computer usable volatile memory 106, e.g. random access memory (RAM), coupled to bus 102 for storing information and instructions for central processor unit 104. System 100 also includes computer usable non-volatile memory 108, e.g. read only memory (ROM), coupled to bus 102 for storing static information and instructions for the central processor unit 104. Such static information is comprised, in one embodiment, of commands for configuration and initial operations of computer system 100. Computer system 100 also includes a data storage unit 110 (e.g., a magnetic or optical disk and disk drive) coupled to bus 102 for storing information and instructions.

System 100 of the present invention also includes an optional alphanumeric input device 112 including alphanumeric and function keys coupled to bus 102 for communicating information and command selections to central processor unit 104. System 100 also optionally includes an optional cursor control device 114 coupled to bus 102 for communicating user input information and command selections to central processor unit 104. System 100 of the present embodiment also includes an optional display device 116 coupled to bus 102 for displaying

information. System 100 of the present embodiment also includes a communication interface 118 which enables computer system 100 to interface with other computers or devices. In one embodiment, communication 118 is, for example, a modem, an integrated services
 5 digital network (ISDN) card or the like, a local area network (LAN) port, etc. Those skilled in the art will recognize that modems or various types of network interface cards (NICs) typically provide data communications via telephone lines, while a LAN port provides data communications via a LAN. Communication interface 118 of computer system 100 may also
 10 enable wireless communications. Furthermore, communication interface 118 may enable communication with other computers or devices through one or more networks. For example, computer system 100, using communication interface 118, may communicate to the "Internet."

15 Computer system 100 may be used to implement the techniques described below. In various embodiments, processor 104 performs the steps of the techniques by executing instructions brought to RAM 106. In alternative embodiments, hard-wired circuitry may be used in place of or in combination with software instructions to implement the described
 20 techniques. Consequently, embodiments of the invention are not limited to any one or a combination of software, hardware, or circuitry.

Instructions executed by processor 104 may be stored in and carried through one or more computer-readable media, which refer to any
 25 medium from which a computer reads information. Computer-readable media may be, for example, a floppy disk, a hard disk, a zip-drive cartridge, a magnetic tape, or any other magnetic medium, a CD-ROM, a CD-RAM, a DVD-ROM, a DVD-RAM, or any other optical medium, paper-tape, punch-cards, or any other physical medium having patterns
 30 of holes, a RAM, a ROM, an EPROM, or any other memory chip or cartridge. Computer-readable media may also be coaxial cables, copper wire, fiber optics, acoustic, or light waves, etc. As an example, the instructions to be executed by processor 104 are in the form of one or more software programs and are initially stored in a CD-ROM being interfaced
 35 with computer system 100. Computer system 100 loads these instructions in RAM 106, executes some instructions, and sends some instructions via communication interface 118, a modem, and a telephone line to a network, the Internet, etc. A remote computer, receiving data through a network cable, executes the received instructions and sends the data to

computer system 100 to be stored in storage device 110.

Referring still to Figure 1, optional display device 116 of Figure 1, may be a liquid crystal device, cathode ray tube, or other display device
 5 suitable for creating graphic images and alphanumeric characters recognizable to a user. Optional cursor control device 114 allows the computer user to dynamically signal the two dimensional movement of a visible symbol (cursor) on a display screen of display device 116. Many implementations of cursor control device 114 are known in the art
 10 including a trackball, mouse, touch pad, joystick or special keys on alphanumeric input device 112 capable of signaling movement of a given direction or manner of displacement. Alternatively, it will be appreciated that a cursor can be directed and/or activated via input from alphanumeric input device 112 using special keys and key sequence
 15 commands. The present invention is also well suited to directing a cursor by other means such as, for example, voice commands. A more detailed discussion of the present invention is found below.

GENERAL METHOD AND SYSTEM 20 FOR INSTRUMENTING A VIRTUAL FUNCTION

With reference next to flow chart 200 of Figure 2 and to Figure 1, exemplary steps used by the various embodiments of present invention are illustrated. Flow chart 200 includes processes of the present invention
 25 which, in one embodiment, are carried out by a processor under the control of computer-readable and computer-executable instructions. The computer-readable and computer-executable instructions reside, for example, in data storage features such as computer usable volatile memory 106, computer usable non-volatile memory 108, and/or data storage device 110 of Figure 1. In one embodiment, the computer-readable
 30 and computer-executable instructions are used to control or operate in conjunction with, for example, processor 104 of Figure 1.

With reference again to Figure 2, steps performed in accordance with one embodiment of the present invention are shown. Although
 35 specific steps are disclosed in flow chart 200 of Figure 2, such steps are exemplary. That is, the present invention is well suited to performing various other steps or variations of the steps recited in Figure 2. At step 202, the present embodiment, determines whether a virtual table exists for a virtual function. More specifically, in one embodiment, an

instrumenting process (instrumentor) monitors a target process using, for example, the debug interface. In this embodiment, the instrumentor monitors the target process for mainly two events: the functions invoked in the target process; and shared modules getting loaded and/or unloaded.

5 Since most compilers maintain virtual table information in the symbol table, in the present embodiment, the instrumentor reads the entire symbol table and knows the exact location of all virtual tables when the module is loaded.

10 Referring still to step 202, in one approach, the instrumenting application (i.e. the instrumentor) is comprised of the Caliper application by Hewlett-Packard Company of Palo Alto, California. The present invention is, however, well suited to use with various other instrumenting applications. Also, in one embodiment, processor 104 of Figure 1 in
15 conjunction with instructions, residing, for example, in RAM 106, ROM 108, and/or data storage device 110, comprise an apparatus which operates to perform step 202.

With reference next to step 204, the present embodiment determines
20 the call type for the virtual function. That is, by locating a virtual table, the present embodiment establishes that the virtual function can be called by a virtual function call. If there is no virtual table corresponding to the virtual function, the present embodiment establishes that the virtual function is called by a non-virtual call (i.e. a direct function call or an
25 indirect function call). The present invention is, however, well suited to use with various other instrumenting applications. Also, in one embodiment, processor 104 of Figure 1 in conjunction with instructions, residing, for example, in RAM 106, ROM 108, and/or data storage device 110, comprise an apparatus which operates to perform step 204. Thus, the
30 present embodiment provides a method and system for determining the type of call made to the virtual function.

With reference now to Figure 3, a flow chart 300 of steps performed in accordance with another embodiment of the present invention is
35 shown. The method of the present embodiment includes steps 202 and 204 of Figure 2. These steps were described above in detail in conjunction with the description of Figure 2, and are not repeated herein for purposes of brevity and clarity. The method of the present embodiment as described in Figure 3 also includes new step 302. At step 302, the present embodiment,

performs instrumentation on said virtual function based upon the call type determined in steps 202 and 204. In one embodiment, if the call type is determined to be non-virtual (i.e. a direct function call or an indirect function call), the present embodiment performs a conventional
 5 instrumentation process upon the virtual function being called. On the other hand, if the call to the virtual function is determined to be a virtual function call, the present embodiment performs a novel instrumentation process. Various embodiments of the novel instrumentation process performed by the present invention are described below in detail. In one
 10 embodiment, processor 104 of Figure 1 in conjunction with instructions, residing, for example, in RAM 106, ROM 108, and/or data storage device 110, comprise an apparatus which operates to perform step 302.

With reference now to Figure 4, a flow chart 400 of steps performed
 15 in accordance with another embodiment of the present invention is shown. The method of the present embodiment includes steps 202, 204 and 302 of Figure 3. These steps were described above in detail in conjunction with the description of Figures 2 and 3, and are not repeated herein for purposes of brevity and clarity. The method of the present
 20 embodiment as described in Figure 4 also includes new steps 402, 404 and 406. At step 402, provided that a virtual table was located for the virtual function, the present embodiment replaces an existing address for the virtual function with a new address in the virtual table such that the new address points to instrumentation code. Also, the present invention is
 25 well suited to an embodiment in which the instrumentation code is dynamically created, and to an embodiment in which the instrumentation code is not dynamically created.

Still referring to step 402, in one embodiment, using the starting
 30 address of each virtual table, the instrumentor loads the data segment of the executable, which is where the virtual tables are stored in one architecture, and reads the corresponding location in the data segment to determine which functions are contained inside each located virtual table. In one embodiment, because shared libraries are relocated to new and
 35 varying addresses on loading of the library, the present invention performs some pointer manipulation to get the virtual address from the address stored in the data segment of the shared library. In such an approach, this offset (i.e. the pointer manipulation) is stored by the instrumentor during the loading of the library.

At step 402, the instrumentor provides instrumentation code. More specifically, in one embodiment, the instrumentor creates for each virtual function found in the virtual tables a single bundle or instruction in
5 shared memory that only contains a breakpoint instruction or other control transfer instruction. Additional details related to instrumentation including discussion of features such as breakpoints or other control transfer instructions, branches, switch tables, procedure lookup tables (PLTs) can be found in co-owned, commonly-assigned U.S. Patent
10 Application Serial No. 09/833,248 filed April 11, 2001, entitled "Dynamic Instrumentation Of An Executable Program", to Hundt et al. which is incorporated herein by reference as background material.

With reference still to step 402, the instrumentor of the present
15 embodiment stores the virtual table information read from the data segment. The instrumentor then rewrites the virtual tables of the target process with a modified virtual table which contains the addresses of the breakpoints associated with each virtual function instead of the original virtual function start address. In one approach, the instrumentation code
20 is comprised of an instruction set which will surrender control of the target process to the instrumentor.

Referring next to step 404, upon a call to the virtual function, in the present embodiment, the target process loads the new address from the
25 virtual table such that execution is directed to the instrumentation code. That is, when the target process makes a virtual function call, it reads the contents of the virtual table to get the function start address. In this embodiment, instead of making an indirect call to the virtual function start address, the target process makes an indirect call to the breakpoint
30 (created at step 402) which is intended to deliver control of the target process to the instrumentor.

With reference now to step 406, in one embodiment, the present invention then continues execution and executes the instrumentation code
35 such that control is delivered to the instrumentor. In one embodiment, processor 104 of Figure 1 in conjunction with instructions, residing, for example, in RAM 106, ROM 108, and/or data storage device 110, comprise an apparatus which operates to perform steps 402, 404, and 406.

With reference now to Figure 5, a flow chart 500 of steps performed in accordance with another embodiment of the present invention is shown. The method of the present embodiment includes steps 202, 204, 302, 402, 404, and 406 of Figure 4. These steps were described above in detail in conjunction with the description of Figure 4, and are not repeated herein for purposes of brevity and clarity. The method of the present embodiment as described in Figure 5 also includes new steps 502 and 504. At step 502, the present embodiment performs a desired task (e.g. an instrumentation task) by the instrumentor. As an example, the present invention enables the instrumentor to, for example, increment a variable representing the virtual function. Although such specific examples are provided herein, the present invention is also well suited to having the instrumentor perform various other tasks.

With reference now to step 504, the present embodiment then resumes execution by the instrumentor at the existing address previously contained in the virtual table. That is, in this embodiment, the instrumentor, without doing any further modification to the target process, continues execution at the original virtual function entry point. In one embodiment, processor 104 of Figure 1 in conjunction with instructions, residing, for example, in RAM 106, ROM 108, and/or data storage device 110, comprise an apparatus which operates to perform steps 502 and 504.

With reference now to Figure 6, a flow chart 600 of steps performed in accordance with another embodiment of the present invention is shown. The method of the present embodiment includes steps 202, 204, 302, 402, 404, and 406 of Figure 4. These steps were described above in detail in conjunction with the description of Figure 4, and are not repeated herein for purposes of brevity and clarity. The method of the present embodiment as described in Figure 6 also includes new steps 602 and 604. At step 602, the present embodiment overwrites the instrumentation code with instrumentation code which performs a desired instrumentation task. That is, in the embodiment of Figure 5, the instrumentor performs the desired instrumentation task and then the instrumentor resumes execution at the existing address previously contained in the virtual table. In the present embodiment however, instrumentation code (e.g. probe code) is added to the target process in the form of instructions which are written over the instrumentation code. The instrumentation code, when

executed by the target process performs the desired instrumentation task. In one embodiment, the instrumentation task comprises incrementing a variable representing the virtual function. Although such a specific example is provided herein, the present invention is also well suited to providing instrumentation code which, when executed, performs various other instrumentation tasks.

With reference still to step 602, in another embodiment, the instrumentation code is provided at a new location. In one such embodiment, the instrumentor must then overwrite all virtual table entries associated with the virtual function to contain the address of the instrumentation code instead of the breakpoint.

At step 604, the present embodiment provides an instruction at the end of the instrumentation code wherein the instruction points back to the existing address previously contained in the virtual table or the address of an instrumented version of the function. In one example, the instruction at the end of the instrumentation code is comprised of a direct branch at the end of it to the virtual function start address. Thus, the present embodiment does not surrender or repeatedly transfer control of the target process to the instrumentor. In one embodiment, processor 104 of Figure 1 in conjunction with instructions, residing, for example, in RAM 106, ROM 108, and/or data storage device 110, comprise an apparatus which operates to perform steps 602 and 604.

With reference now to Figure 7, a flow chart 700 of steps performed in accordance with another embodiment of the present invention is shown. The method of the present embodiment includes steps 202 and 204 of Figure 7. These steps were described above in detail in conjunction with the description of Figure 7, and are not repeated herein for purposes of brevity and clarity. The method of the present embodiment as described in Figure 7 also includes new step 702. At step 702, the present embodiment As yet another example, the present invention is also well suited to an embodiment which determines from which location the virtual function has been called. More specifically, in one embodiment, the instrumentor determines the location from which the virtual function has been called by examining the return pointer. Hence, the present embodiment is able to provide such valuable information to the programmer. In one embodiment, processor 104 of Figure 1 in conjunction with instructions,

residing, for example, in RAM 106, ROM 108, and/or data storage device 110, comprise an apparatus which operates to perform step 702.

With reference now to Figure 8, a flow chart 800 of steps performed in accordance with another embodiment of the present invention is shown. The method of the present embodiment includes steps 202, 204, 302, 402, 404, and 406 of Figure 4. These steps were described above in detail in conjunction with the description of Figure 4, and are not repeated herein for purposes of brevity and clarity. The method of the present embodiment as described in Figure 8 also includes new steps 802. At step 802, the present embodiment maintains a mapping between the existing address for the virtual function and the new address for the virtual function. In so doing, the present embodiment enables the functionality recited above in step 402. That is, by maintaining a mapping between the existing address for the virtual function and the new address for the virtual function, the present embodiment enables the replacing of the existing address of the virtual function with the new address of the virtual function and also replacing of the new address of the virtual function with the existing address of the virtual function.

Hence, the present invention provides in various embodiments, accurate analysis of the number and type of virtual function calls and their callers. Such information is extremely valuable to software developers using object-oriented languages. Since virtual function calls are slower than direct function calls, by presenting this information to the developer the present invention helps gauge whether the software should be altered to change the virtual function call to a non-virtual function call. In an instrumentation tool which relocates the instrumented function into shared memory, various embodiments of the present invention achieve an actual performance improvement because without rewriting the virtual tables for the new function location, a virtual function call is made to the start address of the original function. The original function is then structured to make a long branch to the start of the instrumented function in shared memory. Hence, only one (i.e. the first) virtual function call is made in such embodiments.

Thus, the present invention provides a method and system for instrumenting a virtual function including determining the type of call made to the virtual function.

The foregoing descriptions of specific embodiments of the present invention have been presented for purposes of illustration and description. They are not intended to be exhaustive or to limit the invention to the precise forms disclosed, and obviously many modifications and variations are possible in light of the above teaching. The embodiments were chosen and described in order to best explain the principles of the invention and its practical application, to thereby enable others skilled in the art to best utilize the invention and various embodiments with various modifications as are suited to the particular use contemplated. It is intended that the scope of the invention be defined by the Claims appended hereto and their equivalents.